UNITED STATES DEPARTMENT OF COMMERCE
Patent and Trademark Office
Address: COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

C

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.
09/684,725	10/06/00	HARLAND	L PCS10361ADAM

EXAMINER

HM12/1102

GREGG C BENSON
PFIZER INC
PATENT DEPARTMENT MS
4159 EASTERN POINT ROAD
GROTON CT 06340

ART UNIT	PAPER NUMBER
1646	8

DATE MAILED: 11/02/01

Please find below and/or attached an Office communication concerning this application or proceeding.

Commissioner of Patents and Trademarks

Office Action Summary	Application No. 09/684,725	Applicant(s) HARLAND, LEE	
	Examiner Ruixiang Li	Art Unit 1646	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☐ Responsive to communication(s) filed on _____ .

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-12 and 22* is/are pending in the application.

4a) Of the above claim(s) *13-21* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-12 and 22* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☒ Claim(s) *1-22* are subject to restriction and/or election requirement.

Application Papers

9)☒ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on *10/06/2000* is/are: a)☒ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

11)☐ The proposed drawing correction filed on _____ is: a)☐ approved b)☐ disapproved by the Examiner.

If approved, corrected drawings are required in reply to this Office action.

12)☐ The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

13)☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☒ Some * c)☐ None of:

1.☒ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____ .

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

14)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).

a) ☐ The translation of the foreign language provisional application has been received.

15)☐ Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .
4) ☐ Interview Summary (PTO-413) Paper No(s). _____ .
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: .

DETAILED ACTION

Election/Restrictions

1. Restriction to one of the following inventions is required under 35 U.S.C. 121:

   I. Claims 1-12, and 22, drawn to polynucleotides, primers, vectors, host cells, and the process of making the polypeptide, classified in class 536, subclasses 23.1 and 24.33; class 435, subclasses 320.1, 325, and 69.1.

   II. Claim 13, drawn to polypeptides, classified in class 530, subclass 324.

   III. Claims 14 and 16, drawn to antibodies, classified in class 530, subclass 387.9.

   IV. Claims 15 and 17, drawn to polypeptide modulators, classified in class 530, subclass 300.

   V. Claim 18, drawn to a method of identifying peptide-binding compounds, classified in class 435, subclass 7.1.

   VI. Claim 19, drawn to a method of treatment with antibodies, classified in class 424, subclass 134.1.

   VII. Claims 20 and 21, drawn to a method of treatment with peptide modulators, classified in class 514, subclass 2.

2. The inventions are distinct, each from the other for the following reasons. Inventions I, II, III, and IV are unrelated. Inventions are unrelated if it can be shown that they are not disclosed as capable of use together and they have different modes of operation, different functions, or different effects (MPEP §806.04, MPEP §808.01). In the instance case, the different inventions are drawn to completely different

products having completely different structures and biological functions which are not interchangeable and which require non-cohesive searches and considerations.

3. Inventions V, VI and VII are unrelated. Inventions are unrelated if it can be shown that they are not disclosed as capable of use together and they have different modes of operation, different functions, or different effects (MPEP §806.04, MPEP §808.01). In the instance case the different inventions are drawn to completely different methods each having completely different method steps, using different compositions, and having completely different outcomes. The method of identifying peptide-binding compounds will not provide information regarding the method of treatment with either an antibody or a peptide modulator; On the other hand, the method of treatment with either an antibody or a peptide modulator will not be able to identify the peptide-binding compounds and the three methods are exclusive.
4. Inventions II and V are related as product and process of use. The inventions can be shown to be distinct if either or both of the following can be shown: (1) the process for using the product as claimed can be practiced with another materially different product or (2) the product as claimed can be used in a materially different process of using that product (MPEP § 806.05(h)). In the instant case, the polypeptide may be used in materially different methods, such as production of antibody by immunization of the mice.
5. Inventions III and VI are related as product and process of use. The inventions can be shown to be distinct if either or both of the following can be shown: (1) the process for using the product as claimed can be practiced with another materially

different product or (2) the product as claimed can be used in a materially different process of using that product (MPEP § 806.05(h)). In the instant case, the antibodies may be used to detect or isolate the peptide.

6. Inventions IV and VII are related as product and process of use. The inventions can be shown to be distinct if either or both of the following can be shown: (1) the process for using the product as claimed can be practiced with another materially different product or (2) the product as claimed can be used in a materially different process of using that product (MPEP § 806.05(h)). In the instant case the peptide modulators can be used in a binding assay or to study the biological functions of the peptide.
7. Invention I is an independent invention from V, VI, and VII; Invention II is an independent invention from VI and VII; Invention III is an independent invention from V and VII; Invention IV is an independent invention from V and VI. The different inventions are drawn to distinct product and method inventions.
8. Because these inventions are distinct for the reasons given above and have acquired a separate status in the art as shown by their different classification, restriction for examination purposes as indicated is proper.
9. Because these inventions are distinct for the reasons given above and have acquired a separate status in the art because of their recognized divergent subject matter, restriction for examination purposes as indicated is proper.

10. Because these inventions are distinct for the reasons given above and the search required for a single group is not required for any other group, restriction for examination purposes as indicated is proper.

11. During a telephone conversation with Deborah A. Martin on August 9, a provisional election was made without traverse to prosecute the invention of group I, claims 1-12 and 22. Affirmation of this election must be made by applicant in replying to this Office action. Claims 13-21 are withdrawn from further consideration by the examiner, 37 CFR 1.142(b), as being drawn to a non-elected invention.

Applicant is reminded that upon the cancellation of claims to a non-elected invention, the inventorship must be amended in compliance with 37 CFR 1.48 (b) if one or more of the currently named inventors is no longer an inventor of at least one claim remaining in the application. Any amendment of inventorship must be accompanied by a petition under 37 CFR 1.48 (b) and by the fee required under 37 CFR 1.17 (I).

Minor Objection to the Application

12. The title of the invention is not descriptive. A new title is required that is clearly indicative of the invention to which the claims are directed. The following title is suggested: "A Subtype of Neuromedin U Receptor".

Claim Rejections—35 USC § 101

13. 35 U.S.C. 101 reads as follows:

> Whoever invents or discovers any new and useful process, machine, manufacture, or composition of matter, or any new and useful improvement thereof, may obtain a patent therefor, subject to the conditions and requirements of this title.

14. Claims 1-12 and 22 are rejected under 35 U.S.C. 101 because the claimed invention is not supported by either a specific and substantial asserted utility or a well-established utility.

Claims 1-12 and 22 are drawn to polynucleotides encoding G-protein-coupled receptors and methods of expressing these proteins. The claimed polynucleotides are not supported by either a specific and substantial asserted utility or a well-established utility. A specific and substantial utility is one that is particular to the subject matter claimed and that identifies a "real world" context of use for the claimed invention which does not requires further research. The asserted utility of use of compounds (agonists or antagonists of the claimed polypeptide) in the manufacture of a medicament for treatment of obesity (page 8, line 36-page 8, line 11), while specific, is not substantial because it would require or constitute carrying out further research to identify or reasonably confirm a "real world" context of use. See *Brenner v. Manson*, 383 U.S. 519, 148 USPQ 689 (Sup. Ct. 1966), noting that "a patent is not a hunting license. It is not a reward for the search, but compensation for its successful conclusion."

Likewise, the following asserted utility (page 9, lines 15-22) is not specific and substantial, "Drugs that modulate the novel PFI-002 receptor will therefore be likely to modulate signal transduction processes. It is therefore likely that modulators of the PFI-002 receptor may be useful for the treatment of many different disorders associated with signal transduction that will most likely include,

but not limited to, obesity, diabetes...". The word "likely" and the disclosure of a list of disorders in the specification clearly indicate that the asserted utility requires further research to determine the specific role of the claimed novel PFI-002 receptor in signal transduction processes and its association with a specific disorder.

In addition, the asserted utility of the PFI-002 polypeptide and nucleotide sequences encoding the polypeptide for screening drug candidates for treatment of diseases (page 9, lines 24-35) is not specific and substantial, either. Furthermore, the specification asserts that nucleotide sequences encoding a PFI-002 receptor provide probes to detect chromosomal aberration (page 19, lines 15-21). Since a disease specifically associated with an abnormal level of nucleotide sequences encoding a PFI-002 receptor has not been identified, the use of the probe to diagnose such chromosomal aberration would require further research. Thus, the asserted utility is not specific and substantial.

The invention also lacks a well-established utility. A well-established utility is a specific, substantial, and creditable utility that is well known, immediately apparent, or implied by the specification's disclosure of the properties of a material. Neither the specification as filed nor any art of record before the priority date of this application discloses or suggests any property or activity for the PFI-002 polypeptide and/or nucleotide sequences encoding the polypeptide such that another non-asserted utility would be well-established for the compounds.

15. Claims 1-12 and 22 are also rejected under 35 U.S.C. 112, first paragraph. Specifically, since the claimed invention is not supported by either a specific and

substantial asserted utility or a well established utility for the reasons set forth above, one skilled in the art clearly would not know how to use the claimed invention.

Further, even if the polynucleotide of SEQ ID NO: 1 or encoding a polypeptide of SEQ ID NO: 2 were to have a patentable use, the instant disclosure would not be found to be enabling for the claimed genus of homologues or fragments of the polynucleotides.

The factors to be considered when determining whether a disclosure satisfies enablement requirement include: (i) the quantity of experimentation necessary; (ii) the amount of direction or guidance presented; (iii) the existence of working examples; (iv) the nature of the invention; (v) the state of the prior art; (vi) the relative skill of those in the art; (vii) the predictability or unpredictability of the art; and (viii) the breadth of the claims. *Ex Parte Forman*, 230 USPQ 546 (Bd Pat. App. & Int. 1986); *In re Wands*, 858 F. 2d 731, 8 USPQ 2d 1400 (Fed. Cir. 1988).

Each of the claims recites a genus of polynucleotides that has at least 70% (75%, 80%, 85%, 90%, or 95%) identity to a polynucleotide comprising SEQ ID NO: 1 or encoding a polypeptide of SEQ ID NO: 2, as well as polynucleotides comprising fragments of a polynucleotide of SEQ ID NO: 1 or encoding SEQ ID NO: 2. However, other than SEQ ID NOS: 1 and 2, the disclosure does not provide sufficient guidance and information regarding the structural and functional requirements commensurate in scope with what is encompassed by the instant claims. The disclosure has not shown what modifications (e.g., substitutions,

deletions or additions) one can make to SEQ ID NO: 1 or SEQ ID NO: 2 will result in mutants with the same function as SEQ ID NO: 1 or SEQ ID NO: 2. The state of the art (See, e.g., Ngo, et al, *The Protein Folding Problem and Tertiary Structure Prediction*, 1994, Merz, et al. (ed.), Birkhauser, Boston, MA, pp. 433 and 492-495) is such that the sequence of a peptide and its activity is not well understood and is not predictable. Excising out portions of a protein or modifications to a protein, e.g., by substitutions and deletions, would often result in deleterious effects to the overall activity and effectiveness of the protein.

The instant claims also embrace polynucleotide fragments of any size. However, the disclosure has not shown which portions or fragments of SEQ ID NO: 1 or SEQ ID NO: 2 are critical to the activity of the polypeptide encoded by the claimed polynucleotides. Thus, the disclosure has not provided sufficient guidance and information to enable one skilled in the art to predict which if any fragments of the whole molecule would be reasonably expected to retain characteristic activities alone. The general disclosure that one could make and use SEQ ID NO: 1 or SEQ ID NO: 2 could not be used to be such guidance as to guide one of skill in the art to make and use the invention commensurate in scope with the claims.

Accordingly, the disclosure fails to enable such a myriad of the claimed polynucleotides that not only vary substantially in length but also in amino acid composition and to provide any guidance to those skilled generally on how to make and use useful polynucleotides. Thus, it would require undue experimentation for one skilled in the art to make and use the polynucleotides embraced by the instant

claims.

In addition, the disclosure provides the asserted use of four probes represented by SEQ ID NOS 3-6 in isolation of PFI-002 and determination of tissue distribution of PFI-002. However, since Claim 8, as it is written, comprises at least 15 contiguous nucleotides, it encompasses virtually any random sequence of any length as long as it has a stretch of at least 15 consecutive nucleotides that is the same as in the sequence(s) recited in claim 1. The state of the art is such that determining the specificity of hybridization probes is empirical by nature and the effect of mismatches within an oligonucleotide probe or primer is unpredictable. In view of the factors, the empirical and unpredictable nature of the art and the lack of guidance with respect to how to use other probes within the scope of the claim, the disclosure does not teach one skilled in the art how to successfully use probes of the claimed scope without undue experimentation.

Claim Rejections—35 USC § 112, 1st paragraph

16. The following is a quotation of the first paragraph of 35 U.S.C. 112:

> The specification shall contain a written description of the invention, and of the manner and process of making and using it, in such full, clear, concise, and exact terms as to enable any person skilled in the art to which it pertains, or with which it is most nearly connected, to make and use the same and shall set forth the best mode contemplated by the inventor of carrying out his invention.

17. Claims 1-12 and 22 are rejected under 35 U.S.C. 112, first paragraph, as containing subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor, at the time the application was filed, had possession of the claimed invention.

The description discloses a nucleotide sequence set forth in SEQ ID NO: 1, which encodes a polypeptide as set in SEQ ID NO: 2. However, each of these claims as written includes a genus of polynucleotides that has at least 70% (75%, 80%, 85%, 90%, or 95%) identity to a polynucleotide comprising SEQ ID NO 1 or encoding SEQ ID NO 2, or a polynucleotide encoding the polypeptide expressed by the DNA contained in the clone deposited as NCIMB 41066. Thus, the claims encompass a huge number of polynucleotides that vary substantially both in length and in amino acid composition.

The instant disclosure of a single species of nucleic acid of SEQ ID NO: 1 encoding a single polypeptide of SEQ ID NO: 2 does not adequately support the scope of the claimed genus, which encompasses a substantial variety of subgenera including full-length genes. A description of a genus of cDNA may be achieved by means of a recitation of a representative number of cDNAs, defined by nucleotide sequence, falling within the scope of the genus, or of a recitation of structural features common to the genus, which features constitute a substantial portion of the genus. *Regents of the University of California v. Eli Lilly & Co.*, 119 F3d 1559, 1569, 43 USPQ2d 1398, 1406 (Fed. Cir. 1997). The instant disclosure fails to provide sufficient description information, such as definitive structural or functional features of the claimed genus of polynucleotides. There is no description of the conserved regions that are critical to the structure and function of the genus claimed. There is no description of the sites at which variability may be tolerated and there is no information regarding the relation of structure to function. Furthermore, the prior art

does not provide compensatory structural or correlative teachings to enable one skilled in the art to identify the encompassed polynucleotides as being identical to those instantly claimed.

Due to the breadth of the claim genus and lack of the definitive structural or functional features of the claimed genus, one skilled in the art would not recognize from the disclosure that the applicant was in possession of the claimed genus.

Claim Rejections—35 USC § 112, 2nd paragraph

18. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

19. Claims 1, 8, and 22 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

Claim 1 is indefinite because it recites an abbreviated term, "NCIMB", which should be spelled out in the claim. In addition, the claim can be more clearly stated as "a polynucleotide encoding the polypeptide expressed by the DNA contained in the clone deposited as National Collections of Industrial and Marine Bacteria Limited (NCIMB) 41066";

Claim 8 is vague and indefinite because it is not clear which polynucleotide of claim 1 is referred to by "said polynucleotide".

Claim 22 is indefinite because it uses the words" and/or" to describe an animal cell genetically modified to increase or decrease the expression of a polynucleotide sequence.

20. Claim 22 is also objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim. Applicant is required to cancel the claim(s), or amend the claim(s) to place the claim(s) in proper dependent form, or rewrite the claim(s) in independent form. Claim 22 depends on an unelected claim (Claim 13). For prosecution purpose, the examiner considers as if it depends upon Claim 1.

Claim Rejections—35 USC § 102

21. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

> A person shall be entitled to a patent unless –
>
> (b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

22. Claims 1-10 and 22 are rejected under 35 U.S.C. 102(b) as being anticipated by Database EMBL Accession No. AC008571 (August 3, 1999), which discloses a nucleotide sequence comprising SEQ ID NO: 1 (See attached sequence alignment) as well as a nucleotide sequence capable of hybridizing to the nucleotide sequence. The nucleotide sequence is present in a vector which in turn is present in a host cell, meeting the limitation of claims 9, 10, and 22.

23. Claim 1 is also rejected under 35 U.S.C. 102(b) as being anticipated by Lubert Stryer (*Biochemistry*, 3rd Edition, pp71-90, W. H. Freeman, 1988). Stryer teaches the structure of nucleotides, DNA, and RNA. Claim 1 recites "a polynucleotide fragment" which encompasses just about everything, from a single nucleotide to at least SEQ ID NO: 1 or its complement. Thus, the reference by Stryer anticipates Claim 1.

24. Claims 1, 7-12, and 22 are rejected under 35 U.S.C. 102(b) as being anticipated by Tan, et al (IDS, paper #7). Tan et al disclose a nucleotide sequence of a human G-protein-coupled receptor (FM3) with two regions that contain 19 and 26 contiguous nucleotides identical to that set forth in SEQ ID NO: 1 (See attached sequence alignment). Also disclosed are a nucleotide sequence of mouse G-protein-coupled receptor (FM3) with 15 contiguous nucleotides identical to that set forth in SEQ ID NO: 1 (See attached sequence alignment) and the use of a radiolabled probe encompassing mouse FM-3 ORF in determining the expression profile of FM3 mRNA in several mouse tissues (See Fig. 3, page 228). Furthermore, Tan, et al disclose transfected HEK-293 cells expressing FM3 (cell membrane expression of FM3; See page 228, column 1, lines 24-27), which meets the limitation of claims 10-12 and 22.

25. Claim 12 is also rejected under 35 U.S.C. 102(b) as being anticipated by Maniatis et al (*Molecular Cloning: A laboratory Manual*. 2nd Edition, Book 3, pp17.37-17.41, Cold Spring Harbor Laboratory Press, 1989). Maniatis et al. teach membrane preparation of a cell by centrifugation and thus the reference meets the limitation of

claim 12, which encompasses membrane preparations indistinguishable from those of the art.

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Ruixiang Li whose telephone number is (703) 306-0282. The examiner can normally be reached on Monday through Friday from 8:30 am to 5:00 pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Yvonne Eyler, can be reached on (703) 308-6564. The fax phone number for this Group is (703) 305-3014 or (703) 308-4242.

Communications via Internet e-mail regarding this application, other than those under 35 U.S.C. 132 or which otherwise require a signature, may be used by the applicant and should be addressed to [**yvonne.eyler@uspto.gov**].

All Internet e-mail communications will be made of record in the application file. PTO employees do not engage in Internet communications where there exists a possibility that sensitive information could be identified or exchanged unless the record includes a properly signed express waiver of the confidentiality requirements of 35 U.S.C. 122. This is more clearly set forth in the Interim Internet Usage Policy published in the Official Gazette of the Patent and Trademark on February 25, 1997 at 1195 OG 89.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the Group receptionist whose telephone number is (703) 308-0196.

Ruixiang Li
Examiner
October 16, 2001

Yvonne Eyler
YVONNE EYLER, PH.D
SUPERVISORY PATENT EXAMINER
TECHNOLOGY CENTER 1600

	9	280.8	38.5	1318	89	AF272362	AF272362 Homo sapi
	10	280.8	38.5	187451	64	AC017104	AC017104 Homo sapi
	11	255	35.0	1209	94	AF242873	AF242873 Rattus no
	12	255	35.0	1239	94	AB038649	AB038649 Rattus no
	13	243	33.3	1218	94	AF044602	AF044602 Mus muscu
c	14	159.6	21.9	75950	74	AC073449	AC073449 Homo sapi
	15	134.8	18.5	2040	88	AF034632	AF034632 Homo sapi
	16	134.8	18.5	163284	89	AL137000	AL137000 Human DNA
	17	133.2	18.3	1676	8	AF082210	AF082210 Sphoeroid
	18	127	17.4	1254	45	E11480	E11480 cDNA encodi
	19	127	17.4	4131	93	HSNEURA	X70070 H.sapiens m
	20	120.4	16.5	870	97	HSU60181	U60181 Human growt
	21	120.4	16.5	1101	97	HSU60179	U60179 Human growt
c	22	120.4	16.5	146442	74	AC069523	AC069523 Homo sapi
	23	119.8	16.4	121652	60	AC008191	AC008191 Drosophil
	24	119.8	16.4	140838	65	AC018176	AC018176 Drosophil
	25	119.8	16.4	219832	60	AC007441	AC007441 Drosophil
c	26	119.8	16.4	225374	5	AE003703	AE003703 Drosophil
	27	118.8	16.3	150566	64	AC016938	AC016938 Homo sapi
	28	114	15.6	870	7	SSU60180	U60180 Sus scrofa
	29	114	15.6	1101	7	SSU60178	U60178 Sus scrofa
	30	112.4	15.4	1161	10	I15508	I15508 Sequence 2
	31	112.4	15.4	1161	10	I89330	I89330 Sequence 2
	32	112.4	15.4	1370	10	I20930	I20930 Sequence 1
	33	112.4	15.4	1370	10	I28195	I28195 Sequence 17
	34	112.4	15.4	1466	10	I20931	I20931 Sequence 3
	35	112.4	15.4	1466	10	I28196	I28196 Sequence 19
	36	112.4	15.4	1504	97	HUMD4C	L12398 Homo sapien
	37	112.4	15.4	1610	10	I20932	I20932 Sequence 5
	38	112.4	15.4	1610	10	I28197	I28197 Sequence 21
	39	110	15.1	1529	95	RNNTR2REC	X97121 R.norvegicu
	40	109.2	15.0	1350	94	AB001982	AB001982 Rattus no
	41	109.2	15.0	3129	95	RNU94321	U94321 Rattus norv
	42	106.6	14.6	3917	94	AB017027	AB017027 Mus muscu
	43	101.6	13.9	2379	7	BTD2DOR	X51657 B.taurus do
	44	101	13.9	1089	94	AF149717	AF149717 Rattus no
	45	101	13.9	2351	94	AB015645	AB015645 Rattus no

ALIGNMENTS

```
RESULT   1
AC008571/c
LOCUS       AC008571   216673 bp    DNA             HTG       05-MAY-2000
DEFINITION  Homo sapiens chromosome 5 clone CTC-550M4, WORKING DRAFT SEQUENCE,
            9 unordered pieces.
ACCESSION   AC008571
VERSION     AC008571.3  GI:7708957
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 216673)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 5
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 216673)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On May 5, 2000 this sequence version replaced gi:7211884.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 396672, H361
            Center clone name: CIT-HSPC_550M4
            -------
            Summary Statistics
            Consensus quality: 197198 bases at least Q40
            Consensus quality: 207925 bases at least Q30
            Consensus quality: 210803 bases at least Q20
            Estimated insert size: 223000; pulse field gel estimation
            Estimated insert size: 215873; sum-of-contigs estimation
            Quality coverage:  5.46 in Q20 bases; pulse field gel estimation
            Quality coverage:  5.64 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 9 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     3048: contig of 3048 bp in length
            *     3049     3148: gap of unknown length
            *     3149     8405: contig of 5257 bp in length
            *     8406     8505: gap of unknown length
            *     8506    23492: contig of 14987 bp in length
            *    23493    23592: gap of unknown length
            *    23593    40018: contig of 16426 bp in length
            *    40019    40118: gap of unknown length
            *    40119    54776: contig of 14658 bp in length
            *    54777    54876: gap of unknown length
            *    54877    80050: contig of 25174 bp in length
            *    80051    80150: gap of unknown length
            *    80151   121104: contig of 40954 bp in length
            *   121105   121204: gap of unknown length
            *   121205   167091: contig of 45887 bp in length
            *   167092   167191: gap of unknown length
            *   167192   216673: contig of 49482 bp in length.
FEATURES             Location/Qualifiers
     source          1. .216673
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="5"
                     /clone="CTC-550M4"
                     /clone_lib="CalTech human BAC library C"
BASE COUNT    66753 a  41883 c  42226 g  65000 t    811 others
ORIGIN
```

```
  Query Match             100.0%;  Score 729;  DB 60;  Length 216673;
  Best Local Similarity   100.0%;  Pred. No. 4e-130;
  Matches  729;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 atggaaaaacttcagaatgcttcctggatctaccagcagaaactagaagatccattccag 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84403 ATGGAAAAACTTCAGAATGCTTCCTGGATCTACCAGCAGAAACTAGAAGATCCATTCCAG 84344

Qy      61 aaacacctgaacagcaccgaggagtatctggccttcctctgcggacctcggcgcagccac 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84343 AAACACCTGAACAGCACCGAGGAGTATCTGGCCTTCCTCTGCGGACCTCGGCGCAGCCAC 84284

Qy     121 ttcttcctccccgtgtctgtggtgtatgtgccaatttttgtggtgggggtcattggcaat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84283 TTCTTCCTCCCCGTGTCTGTGGTGTATGTGCCAATTTTTGTGGTGGGGGTCATTGGCAAT 84224

Qy     181 gtcctggtgtgcctggtgattctgcagcaccaggctatgaagacgcccaccaactactac 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84223 GTCCTGGTGTGCCTGGTGATTCTGCAGCACCAGGCTATGAAGACGCCCACCAACTACTAC 84164

Qy     241 ctcttcagcctggcggtctctgacctcctggtcctgctccttggaatgcccctggaggtc 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84163 CTCTTCAGCCTGGCGGTCTCTGACCTCCTGGTCCTGCTCCTTGGAATGCCCCTGGAGGTC 84104

Qy     301 tatgagatgtggcgcaactacccttttcttgttcgggcccgtgggctgctacttcaagacg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   84103 TATGAGATGTGGCGCAACTACCCTTTCTTGTTCGGGCCCGTGGGCTGCTACTTCAAGACG 84044

Qy     361 gccctctttgagaccgtgtgcttcgcctccatcctcagcatcaccaccgtcagcgtggag 420
```

```
Db  84043 GCCCTCTTTGAGACCGTGTGCTTCGCCTCCATCCTCAGCATCACCACCGTCAGCGTGGAG 83984

Qy    421 cgctacgtggccatcctacacccgttccgcgccaaactgcagagcacccggcgccgggcc 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83983 CGCTACGTGGCCATCCTACACCCGTTCCGCGCCAAACTGCAGAGCACCCGGCGCCGGGCC 83924

Qy    481 ctcaggatcctcggcatcgtctggggcttctccgtgctcttctccctgcccaacaccagc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83923 CTCAGGATCCTCGGCATCGTCTGGGGCTTCTCCGTGCTCTTCTCCCTGCCCAACACCAGC 83864

Qy    541 atccatggcatcaagttccactacttccccaatgggtccctggtcccaggttcggccacc 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83863 ATCCATGGCATCAAGTTCCACTACTTCCCCAATGGGTCCCTGGTCCCAGGTTCGGCCACC 83804

Qy    601 tgtacggtcatcaagcccatgtggatctacaatttcatcatccaggtcacctccttccta 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
      83803 TGTACGGTCATCAAGCCCATGTGGATCTACAATTTCATCATCCAGGTCACCTCCTTCCTA 83744

        661 ttctacctcctccccatgactgtcatcagtgtcctctactacctcatggcactcagagtg 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  83743 TTCTACCTCCTCCCCATGACTGTCATCAGTGTCCTCTACTACCTCATGGCACTCAGAGTG 83684

Qy    721 agtatctag 729
          |||||||||
Db  83683 AGTATCTAG 83675
```

RESULT 2
AF242874
LOCUS AF242874 1239 bp mRNA PRI 13-JUL-2000
DEFINITION Homo sapiens neuromedin U receptor 2 (NMU2R) mRNA, complete cds.
ACCESSION AF242874
VERSION AF242874.1 GI:9082155
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1239)
AUTHORS Howard,A.D., Wang,R., Pong,S.S., Mellin,T.N., Strack,A., Guan,X.M., Zeng,Z., Williams,D.L. Jr., Feighner,S.D., Nunes,C.N., Murphy,B., Stair,J.N., Yu,H., Jiang,Q., Clements,M.K., Tan,C.P., McKee,K.K., Hreniuk,D.L., McDonald,T.P., Lynch,K.R., Evans,J.F., Austin,C.P., Caskey,C.T., Van der Ploeg,L.H. and Liu,Q.
TITLE Identification of receptors for neuromedin U and its role in feeding
JOURNAL Nature 406 (6791), 70-74 (2000)
MEDLINE 20351041
REFERENCE 2 (bases 1 to 1239)
AUTHORS Liu,Q., McDonald,T.P., Wang,R., Jiang,Q. and Howard,A.D.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2000) Pharmacology, Merck Research Labs, West Point, PA 19486, USA
FEATURES Location/Qualifiers
source 1..1239
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
gene 1..1239
/gene="NMU2R"
CDS 1..1239
/gene="NMU2R"
/note="FM-4"
/codon_start=1
/product="neuromedin U receptor 2"
/protein_id="AAF82755.1"
/db_xref="GI:9082156"
/translation="MEKLQNASWIYQQKLEDPFQKHLNSTEEYLAFLCGPRRSHFFLP
VSVVYVPIFVVGVIGNVLVCLVILQHQAMKTPTNYYLFSLAVSDLLVLLLGMPLEVYE
MWRNYPFLFGPVGCYFKTALFETVCFASILSITTVSVERYVAILHPFRAKLQSTRRRA
LRILGIVWGFSVLFSLPNTSIHGIKFHYFPNGSLVPGSATCTVIKPMWIYNFIIQVTS
FLFYLLPMTVISVLYYLMALRLKKDKSLEADEGNANIQRPCRKSVNKMLFVLVLVFAI

BA
OR

(

```
GAP of reverse of: us-09-684-725-1 check: 1088  from: 1  to: 729

FROMIG of: /home/bobryen/big/US09684725.seq
sequence 1, application us/09684725
general information:
 applicant:  lee harland
 title of invention:  novel polypeptide
 file reference:  pcs10361adam . . .

to: version1  check: 7472  from: 192001  to: 218807


Symbol comparison table: /opt/gcg/gcgcore/data/rundata/nwsgapdna.cmp
CompCheck: 8760

       Gap Weight:     50      Average Match: 10.000
    Length Weight:      3   Average Mismatch:  0.000

          Quality:   7290             Length:  26807
            Ratio: 10.000               Gaps:      0
Percent Similarity: 100.000   Percent Identity: 100.000

      Match display thresholds for the alignment(s):
                  | = IDENTITY
                  : =   5
                  . =   1

us-09-684-725-1 x version1 September 12, 2001 10:41


   729 .....................................ctagatactcact 717
                                           ||||||||||||
204851 tgaaaaaagaagggatgctgcctcaggccaccccagcctagatactcact 204900

   716 ctgagtgccatgaggtagtagaggacactgatgacagtcatggggaggag 667
       ||||||||||||||||||||||||||||||||||||||||||||||||||
204901 ctgagtgccatgaggtagtagaggacactgatgacagtcatggggaggag 204950

   666 gtagaataggaaggaggtgacctggatgatgaaattgtagatccacatgg 617
       ||||||||||||||||||||||||||||||||||||||||||||||||||
204951 gtagaataggaaggaggtgacctggatgatgaaattgtagatccacatgg 205000

   616 gcttgatgaccgtacaggtggccgaacctgggaccagggacccattgggg 567
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205001 gcttgatgaccgtacaggtggccgaacctgggaccagggacccattgggg 205050

   566 aagtagtggaacttgatgccatggatgctggtgttgggcagggagaagag 517
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205051 aagtagtggaacttgatgccatggatgctggtgttgggcagggagaagag 205100

   516 cacggagaagccccagacgatgccgaggatcctgagggcccggcgccggg 467
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205101 cacggagaagccccagacgatgccgaggatcctgagggcccggcgccggg 205150

   466 tgctctgcagtttggcgcggaacgggtgtaggatggccacgtagcgctcc 417
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205151 tgctctgcagtttggcgcggaacgggtgtaggatggccacgtagcgctcc 205200

   416 acgctgacggtggtgatgctgaggatggaggcgaagcacacggtctcaaa 367
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205201 acgctgacggtggtgatgctgaggatggaggcgaagcacacggtctcaaa 205250

   366 gagggccgtcttgaagtagcagcccacgggcccgaacaagaaagggtagt 317
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205251 gagggccgtcttgaagtagcagcccacgggcccgaacaagaaagggtagt 205300

   316 tgcgccacatctcatagacctccaggggcattccaaggagcaggaccagg 267
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205301 tgcgccacatctcatagacctccaggggcattccaaggagcaggaccagg 205350

   266 aggtcagagaccgccaggctgaagaggtagtagttggtgggcgtcttcat 217
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205351 aggtcagagaccgccaggctgaagaggtagtagttggtgggcgtcttcat 205400

   216 agcctggtgctgcagaatcaccaggcacaccaggacattgccaatgaccc 167
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205401 agcctggtgctgcagaatcaccaggcacaccaggacattgccaatgaccc 205450

   166 ccaccacaaaaattggcacatacaccacagacacggggaggaagaagtgg 117
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205451 ccaccacaaaaattggcacatacaccacagacacggggaggaagaagtgg 205500

   116 ctgcgccgaggtccgcagaggaaggccagatactcctcggtgctgttcag 67
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205501 ctgcgccgaggtccgcagaggaaggccagatactcctcggtgctgttcag 205550

    66 gtgtttctggaatggatcttctagtttctgctggtagatccaggaagcat 17
       ||||||||||||||||||||||||||||||||||||||||||||||||||
205551 gtgtttctggaatggatcttctagtttctgctggtagatccaggaagcat 205600

    16 tctgaagtttttccat................................. 1
       ||||||||||||||||
205601 tctgaagtttttccatccctgacattaaaatccaaggcctgagcctcccc 205650
```

AC 008571

Version AC008571.1

GI : 5686496

(Submitted on August 3, 1999)

```
Rattus.
1  (sites)
Hosoya,M., Moriya,T., Kawamata,Y., Ohkubo,S., Fujii,R., Matsui,H.,
Shintani,Y., Fukusumi,S., Habata,Y., Hinuma,S., Onda,H.,
Nishimura,O. and Fujino,M.
Identification and Functional Characterization of a Novel Subtype
of Neuromedin U Receptor
J. Biol. Chem. 275 (38), 29528-29532 (2000)
10887190
2  (bases 1 to 1314)
Fujii,R., Shintani,Y. and Hinuma,S.
Direct Submission
Submitted (03-APR-2000) to the DDBJ/EMBL/GenBank databases. Shuji
Hinuma, Takeda Chemical Industries, LTD, Discovery Research
Laboratories 1; Wadai 10, Tsukuba, Ibaraki 300-4293, Japan
(E-mail:Hinuma_Shuji@takeda.co.jp, Tel:81-298-64-5035,
Fax:81-298-64-5000)
            Location/Qualifiers
urce        1. .1314
            /organism="Rattus norvegicus"
            /db_xref="taxon:10116"
ne          127. .1314
            /gene="TGR-1"
S           127. .1314
            /gene="TGR-1"
            /codon_start=1
            /product="G protein-coupled receptor TGR-1"
            /protein_id="BAB13722.1"
            /db_xref="GI:10257383"
            /translation="MGKLENASWIHDPLMKYLNSTEEYLAHLCGPKRSDLSLPVSVAY
            ALIFLVGVMGNLLVCMVIVRHQTLKTPTNYYLFSLAVSDLLVLLLGMPLEIYEMWHNY
            PFLFGPVGCYFKTALFETVCFASILSVTTVSVERYVAIVHPFRAKLESTRRRALRILS
            LVWSFSVVFSLPNTSIHGIKFQHFPNGSSVPGSATCTVTKPMWVYNLIIQATSFLFYI
            LPMTLISVLYYLMGLRLKRDESLEANKVAVNIHRPSRKSVTKMLFVLVLVFAICWTPF
            HVDRLFFSFVEEWTESLAAVFNLIHVVSGVFFYLSSAVNPIIYNLLSRRFRAAFRNVV
            SPTCKWCHPRHQPQGPPAQKIIFLTECHLMELTEDAGPQFPGQSSIHNTNLTMAPCAG
            EVP"
OUNT      269 a    397 c    305 g    343 t
```

```
Query Match             64.7%;  Score 471.4;  DB 94;  Length 1314;
Local Similarity   79.8%;  Pred. No. 1.9e-80;
Matches  575;  Conservative    0;  Mismatches  131;  Indels   15;  Gaps    1;

Qy      1 atggaaaaacttcagaatgcttcctggatctaccagcagaaactagaagatccattccag 60
          |||| ||||||| | ||||||||||||||| ||           |||||| ||  |
Db    127 ATGGGAAAACTTGAAAATGCTTCCTGGATCCAC----------GATCCACTCATG 171

Qy     61 aaacacctgaacagcaccgaggagtatctggccttcctctgcggacctcggcgcagccac 120
          ||  || |||||||||| ||||||||  |||||  ||| || |||||   ||||||  ||
Db    172 AAGTACTTGAACAGCACAGAGGAGTACTTGGCCCACCTGTGTGGACCCAAGCGCAGTGAC 231

Qy    121 ttcttcctccccgtgtctgtggtgtatgtgccaatttttgtggtgggggtcattggcaat 180
          | | ||| || |||||||||  |||| ||  || ||  |||||||||| || ||||||
Db    232 CTATCCCTTCCGGTGTCTGTGGCCTATGCGCTGATCTTCCTGGTGGGGGTAATGGGCAAT 291

Qy    181 gtcctggtgtgcctggtgattctgcagcaccaggctatgaagacgcccaccaactactac 240
          | |||||||||  |||||||| | |   || ||| || ||||||| ||||||||||||
Db    292 CTTCTGGTGTGCATGGTGATTGTCCGACATCAGACTTTGAAGACACCCACCAACTACTAT 351

Qy    241 ctcttcagcctggcggtctctgacctcctggtcctgctccttggaatgcccctggaggtc 300
          ||||||||| |||| ||||| || || ||||||||||| | || |||||  |||||  ||
Db    352 CTCTTCAGCTTGGCAGTCTCAGATCTGCTGGTCCTGCTCTTGGGGATGCCTCTGGAAATC 411

Qy    301 tatgagatgtggcgcaactacccttttcttgttcgggcccgtgggctgctacttcaagacg 360
          || |||||||| ||| |||||||| |||||||||| ||||| |||||||||||||||| |||
Db    412 TACGAGATGTGGCACAATTACCCTTTCCTGTTCGGGCCTGTGGGATGCTACTTCAAGACA 471

Qy    361 gccctctttgagaccgtgtgcttcgcctccatcctcagcatcaccaccgtcagcgtggag 420
          |||||||| ||||| |||||||| |||||||| ||||| ||| |||||| ||| |||
Db    472 GCCCTCTTCGAGACTGTGTGCTTTGCCTCCATTCTCAGTGTCACCACGGTTAGCGTAGAG 531

Qy    421 cgctacgtggccatcctacacccgttccgcgccaaactgcagagcacccggcgccgggcc 480
          ||||| |||||||  | ||||| ||||| ||||| ||| ||||||| ||||| ||||||
Db    532 CGCTATGTGGCCATTGTCCACCCTTTCCGAGCCAAGCTGGAGAGCACGCGGCGACGGGCC 591

Qy    481 ctcaggatcctcggcatcgtctggggcttctccgtgctcttctccctgcccaacaccagc 540
          |||||||||||| || | |||||| ||||||| ||| |||| ||  ||||||| ||||||
Db    592 CTCAGGATCCTCAGCCTAGTCTGGAGCTTCTCTGTGGTCTTTTCTTTGCCCAATACCAGC 651

Qy    541 atccatggcatcaagttccactacttccccaatgggtccctggtcccaggttcggccacc 600
          |||||||||||||||||||  |||| ||||| ||||||   || || || || ||||||
Db    652 ATCCATGGCATCAAGTTCCAGCACTTTCCCAACGGGTCCTCCGTACCTGGCTCAGCCACC 711

Qy    601 tgtacggtcatcaagcccatgtggatctacaatttcatcatccaggtcacctccttccta 660
          || || |||| ||| ||||||||| | || || || |||||||| |  |||  ||||||
Db    712 TGCACAGTCACCAAACCCATGTGGGTGTATAACTTGATCATCCAAGCTACCAGCTTCCTC 771

Qy    661 ttctacctcctccccatgactgtcatcagtgtcctctactacctcatggcactcagagtg 720
          |||||| ||||||| ||||| ||||||| ||||||||||||||||||  ||||| ||
Db    772 TTCTACATCCTCCCAATGACCCTCATCAGCGTCCTCTACTACCTCATGGGGCTCAGGCTG 831

Qy    721 a 721
          |
Db    832 A 832
```

```
RESULT   8
HSOGPCR1
LOCUS       HSOGPCR1      828 bp    DNA             PRI       22-JUL-1999
DEFINITION  Homo sapiens orphan G protein-coupled receptor gene, first coding
            exon.
ACCESSION   AF044600
VERSION     AF044600.1  GI:2865467
KEYWORDS    .
SEGMENT     1 of 2
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 828)
  AUTHORS   Tan,C.P., McKee,K.K., Liu,Q., Palyha,O.C., Feighner,S.D.,
            Hreniuk,D.L., Smith,R.G. and Howard,A.D.
  TITLE     Cloning and characterization of a human and murine T-cell orphan
            G-protein-coupled receptor similar to the growth hormone
            secretagogue and neurotensin receptors
  JOURNAL   Genomics 52 (2), 223-229 (1998)
  MEDLINE   99000845
   PUBMED   9782091
REFERENCE   2  (bases 1 to 828)
  AUTHORS   Tan,C.T., McKee,K.K., Palyha,O.C., Feighner,S.D., Hreniuk,D.L.,
            Smith,R.G. and Howard,A.D.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JAN-1998) Biochemistry and Physiology, Merck and Co.
            Inc., P.O. Box 2000, Rahway, NJ 07065, USA
FEATURES             Location/Qualifiers
     source          1. .828
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     exon            <1. .828
                     /note="an intron of about 3kb follows this exon"
BASE COUNT      134 a    267 c    256 g    171 t
ORIGIN
```

```
Query Match             38.5%;  Score 280.8;  DB 93;  Length 828;
Best Local Similarity   65.8%;  Pred. No. 4.6e-44;
Matches  408;  Conservative    0;  Mismatches  212;  Indels    0;  Gaps    0;

Qy    103 ggacctcggcgcagccacttcttcctccccgtgtctgtggtgtatgtgccaatttttgtg 162
          || || | ||  |   |  | ||| | ||| | | ||     ||  |||  || || |||
Db     88 GGGCCCCAGCAGACAGAGCTGTTCATGCCCATCTGTGCCACATACCTGCTGATCTTCGTG 147

Qy    163 gtgggggtcattggcaatgtcctggtgtgcctggtgattctgcagcaccaggctatgaag 222
          |||||  |   | |||||||  |||   || ||||| || ||||  ||| |||| |||
```

```
Qy     48 GTGGGCGCTGTGGGCAATGGGCTGACCTGTCTGGTCATCCTGCGCCACAAGGCCATGCGC 207

Db     23 acgcccaccaactactacctcttcagcctggcggtctctgacctcctggtcctgctcctt 282
          ||||| |||||||||||||||||||||||||| || || ||||| ||||| |||||  |
Db    208 ACGCCTACCAACTACTACCTCTTCAGCCTGGCCGTGTCGGACCTGCTGGTGCTGCTGGTG 267

Qy     83 ggaatgcccctggaggtctatgagatgtggcgcaactacccttttcttgttcgggcccgtg 342
          ||  |||||||||| |||||||||||||| ||||||||| ||| || | ||    |
Db    268 GGCCTGCCCCTGGAGCTCTATGAGATGTGGCACAACTACCCCTTCCTGCTGGGCGTTGGT 327

Qy     43 ggctgctacttcaagacggccctctttgagaccgtgtgcttcgcctccatcctcagcatc 402
          ||||||||  |||   |||   || |||||||  || ||| | |||||  | |||| | ||
Db    328 GGCTGCTATTTCCGCACGCTACTGTTTGAGATGGTCTGCCTGGCCTCAGTGCTCAACGTC 387

Qy     03 accaccgtcagcgtggagcgctacgtggccatcctacacccgttccgcgccaaactgcag 462
          ||  || | |||||||| ||||| |||||| |  | |||||  |||  ||||       |
Db    388 ACTGCCCTGAGCGTGGAACGCTATGTGGCCGTGGTGCACCCACTCCAGGCCAGGTCCATG 447

Qy     63 agcacccggcgccgggccctcaggatcctcggcatcgtctggggcttctccgtgctcttc 522
            || |||  ||  |  | | |  | || ||   |||||||||  |  || ||||||  |
Db    448 GTGACGCGGGCCCATGTGCGCCGAGTGCTTGGGGCCGTCTGGGGTCTTGCCATGCTCTGC 507

Qy     23 tccctgcccaacaccagcatccatggcatcaagttccactacttccccaatgggtccctg 582
          ||||||||||||||||||  | ||  ||||||  |   |   ||  | |||    ||  ||
Db    508 TCCCTGCCCAACACCAGCCTGCACGGCATCCGGCAGCTGCACGTGCCCTGCCGGGGCCCA 567

Qy     83 gtcccaggttcggccacctgtacggtcatcaagcccatgtggatctacaatttcatcatc 642
          ||  ||||  ||  ||    ||  | |  ||   ||   |   |||||||  |  |  |
Db    568 GTGCCAGACTCAGCTGTTTGCATGCTGGTCCGCCCACGGGCCCTCTACAACATGGTAGTG 627

Qy     03 caggtcacctccttcctattctacctcctccccatgactgtcatcagtgtcctctactac 702
          |||  |||| |  | || |||| |  ||| ||||||  |  |||| || || ||||||
Db    628 CAGACCACCGCGCTGCTCTTCTTCTGCCTGCCCATGGCCATCATGAGCGTGCTCTACCTG 687

Qy     03 ctcatggcactcagagtgag 722
          |||||  |  ||  || || |
Db    688 CTCATTGGGCTGCGACTGCG 707
```

AF272362 1318 bp mRNA PRI 23-OCT-2000
Homo sapiens neuromedin U receptor 1 (NMUR1) mRNA, complete cds.
AF272362
AF272362.1 GI:10946200

human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1318)
Raddatz,R., Wilson,A.E., Artymyshyn,R., Bonini,J.A., Borowsky,B., Boteju,L.W., Zhou,S., Kouranova,E.V., Nagorny,R., Guevarra,M.S., Dai,M., Lerman,G.S., Vaysse,P.J., Branchek,T.A., Gerald,C., Forray,C. and Adham,N.
Identification and Characterization of Two Neuromedin U Receptors Differentially Expressed in Peripheral Tissues and the Central Nervous System
J. Biol. Chem. 275 (42), 32452-32459 (2000)
10899166
2 (bases 1 to 1318)
Bonini,J.A., Raddatz,R., Wilson,A. and Borowsky,B.
Direct Submission
Submitted (25-MAY-2000) Target Discovery and Assessment, Synaptic Pharmaceutical Corporation, 215 College Road, Paramus, NJ 07652, USA

```
     Location/Qualifiers
     1. .1318
     /organism="Homo sapiens"
     /db_xref="taxon:9606"
     /chromosome="2"
     /map="2q34-q37"
     1. .1318
     /gene="NMUR1"
CDS  38. .1318
     /gene="NMUR1"
     /note="G protein-coupled receptor"
     /codon_start=1
     /product="neuromedin U receptor 1"
     /protein_id="AAG24793.1"
     /db_xref="GI:10946201"
     /translation="MTPLCLNCSVLPGDLYPGGARNPMACNGSAARGHFDPEDLNLTD
     EALRLKYLGPQQTELFMPICATYLLIFVVGAVGNGLTCLVILRHKAMRTPTNYYLFSL
     AVSDLLVLLVLGLPLELYEMWHNYPFLLGVGGCYFRTLLFEMVCLASVLNVTALSVERY
     VAVVHPLQARSMVTRAHVRRVLGAVWGLAMLCSLPNTSLHGIRQLHVPCRGPVPDSAV
     CMLVRPRALYNMVVQTTALLFFCLPMAIMSVLYLLIGLRLRRERLLLMQEAKGRGSAA
     ARSRYTCRLQQHDRGRRQVTKMLFVLVVVFGICWAPFHADRVMWSVVSQWTDGLHLAF
     QHVHVISGIFFYLGSAANPVLYSLMSSRFRETFQEALCLGACCHRLRPRHSSHSLSRM
     TTGSTLCDVGSLGSWVHPLAGNDGPEAQQETDPS"
BASE COUNT 208 a 441 c 403 g 266 t
ORIGIN
```

```
Query Match          38.5%;  Score 280.8;  DB 89;  Length 1318;
Best Local Similarity  65.8%;  Pred. No. 4.1e-44;
Matches  408;  Conservative    0;  Mismatches  212;  Indels    0;  Gaps    0;

Qy    103 ggacctcggcgcagccacttcttcctccccgtgtctgtggtgtatgtgccaatttttgtg 162
          || || | ||  |   |  | ||| | ||| | | ||     ||  |||  || || |||
Db    194 GGGCCCCAGCAGACAGAGCTGTTCATGCCCATCTGTGCCACATACCTGCTGATCTTCGTG 253

Qy    163 gtgggggtcattggcaatgtcctggtgtgcctggtgattctgcagcaccaggctatgaag 222
          |||||  |   | |||||||  |||   || ||||| || ||||  ||| |||| |||
Db    254 GTGGGCGCTGTGGGCAATGGGCTGACCTGTCTGGTCATCCTGCGCCACAAGGCCATGCGC 313

Qy    223 acgcccaccaactactacctcttcagcctggcggtctctgacctcctggtcctgctcctt 282
          ||||| |||||||||||||||||||||||||| || || |||||  |||||  |||||  |
Db    314 ACGCCTACCAACTACTACCTCTTCAGCCTGGCCGTGTCGGACCTGCTGGTGCTGCTGGTG 373

Qy    283 ggaatgcccctggaggtctatgagatgtggcgcaactacccttttcttgttcgggcccgtg 342
          ||  ||||||||||| |||||||||||||| ||||||||| ||| || | ||    |
Db    374 GGCCTGCCCCTGGAGCTCTATGAGATGTGGCACAACTACCCCTTCCTGCTGGGCGTTGGT 433

Qy    343 ggctgctacttcaagacggccctctttgagaccgtgtgcttcgcctccatcctcagcatc 402
          ||||||||  |||   |||   || |||||||  || ||| | |||||  | |||| | ||
Db    434 GGCTGCTATTTCCGCACGCTACTGTTTGAGATGGTCTGCCTGGCCTCAGTGCTCAACGTC 493

Qy    403 accaccgtcagcgtggagcgctacgtggccatcctacacccgttccgcgccaaactgcag 462
          ||  || | |||||||| ||||| |  | |||||  |||  ||||       |
Db    494 ACTGCCCTGAGCGTGGAACGCTATGTGGCCGTGGTGCACCCACTCCAGGCCAGGTCCATG 553

Qy    463 agcacccggcgccgggccctcaggatcctcggcatcgtctggggcttctccgtgctcttc 522
            || |||  ||  |  | | |  | || ||   |||||||||  |  || ||||||  |
Db    554 GTGACGCGGGCCCATGTGCGCCGAGTGCTTGGGGCCGTCTGGGGTCTTGCCATGCTCTGC 613

Qy    523 tccctgcccaacaccagcatccatggcatcaagttccactacttccccaatgggtccctg 582
          ||||||||||||||||  | ||  ||||||  |   |   ||  | |||    ||  ||
Db    614 TCCCTGCCCAACACCAGCCTGCACGGCATCCGGCAGCTGCACGTGCCCTGCCGGGGCCCA 673

Qy    583 gtcccaggttcggccacctgtacggtcatcaagcccatgtggatctacaatttcatcatc 642
          ||  ||||  ||  ||    ||  | |  ||   ||   |   |||||||  |  |  |
Db    674 GTGCCAGACTCAGCTGTTTGCATGCTGGTCCGCCCACGGGCCCTCTACAACATGGTAGTG 733

Qy    643 caggtcacctccttcctattctacctcctccccatgactgtcatcagtgtcctctactac 702
          |||  |||| |  | || |||| |  ||| ||||||  |  |||| || || ||||||
Db    734 CAGACCACCGCGCTGCTCTTCTTCTGCCTGCCCATGGCCATCATGAGCGTGCTCTACCTG 793

Qy    703 ctcatggcactcagagtgag 722
          |||||  |  ||  || || |
Db    794 CTCATTGGGCTGCGACTGCG 813
```

RESULT 10
AC017104
LOCUS AC017104 187451 bp DNA HTG 16-MAR-2001
DEFINITION Homo sapiens chromosome 2 clone RP11-562I5, WORKING DRAFT SEQUENCE,

A

```
-308  ACCTGCCTGCCTCAGCTTCCTTCGCGTTGGGATTAAAGCTGCGCACTACCACTGCCCGGCCAATTTTATTATTTTCAAGG   -229
-228  TTTGCACTCCGAATACTCTGTAGTTGAAATGCACTTTAGTGGGTGGCAGATGCTTTCTTTCCCAGTGGCATGTGACTAAT   -149
-148  CAGTCCTCTACAGTGTGATAACTTAGGACAGAGCTGGGATTACCTAAATGACTTCTGGGTATCTCCCCCTTCTATCCTAC    -69
 -68  AGACTCCTCCCTGCCTCAATTGTTCCATCTTTCCTGGAGCGCTCTCCCCAAATGCTTCAAGGAGCCCC               -1

+ 1 CTG GTC TGC AAT ATC AGT GAG TTC AAG TGG CCC TAT CAA CCT GAG GAT CTG AAC CTT ACC   60
  1 L   V   C   N   I   S   E   F   K   W   P   Y   Q   P   E   D   L   N   L   T    20
                                                                          TM-1 ——
 61 GAT GAG GCC CTG AGG CTG AAG TAT TTG GGG CCA CAG CAG ATG AAA CAG TTT GTC CCC ATC   120
 21 D   E   A   L   R   L   K   Y   L   G   P   Q   Q   M   K   Q   F   V   P   I    40
 ————————————————————————————————————————————————————————————————————————————

121 TGT GTC ACG TAC CTG CTG ATC TTC GTG GTG GGC ACT CTG GGC AAC GGG CTG ACC TGC ACC   180
 41 C   V   T   Y   L   L   I   F   V   V   G   T   L   G   N   G   L   T   C   T    60
                                                  TM-2————————————————————————
181 GTC ATC CTG CGC AAC AAG ACT ATG CGC ACG CCC ACC AAC TTC TAC CTC TTC AGC CTC GCT   240
 61 V   I   L   R   N   K   T   M   R   T   P   T   N   F   Y   L   F   S   L   A    80
 ————————————————————————————————————————————————————————————————————————————

241 GTG TCC GAT ATG CTG GTG CTC CTG GTG GGC TTG CCT CTG GAG CTT TAT GAG ATG CAG CAA   300
 81 V   S   D   M   L   V   L   L   V   G   L   P   L   E   L   Y   E   M   Q   Q   100
                                                              TM-3————————————————
301 AAT TAC CCG TTC CAG CTG GGT GCG AGT GCC TGC TAC TTC CGA ATA CTG CTC TTA GAG ACC   360
101 N   Y   P   F   Q   L   G   A   S   A   C   Y   F   R   I   L   L   L   E   T   120
 ————————————————————————————————————————————————————————————————————————————

361 GTC TGC CTA GCT TCA GTG CTC AAT GTC ACA GCC CTG AGT GTG GAG CGT TAT GTG GCC GTG   420
121 V   C   L   A   S   V   L   N   V   T   A   L   S   V   E   R   Y   V   A   V   140
                                                                      TM-4————————
421 GTG CGC CCA CTC CAA GCC AAG TCT GTG ATG ACA CGG GCC CAT GTG CGC CGC ATG GTG GGG   480
141 V   R   P   L   Q   A   K   S   V   M   T   R   A   H   V   R   R   M   V   G   160
 ————————————————————————————————————————————————————————————————————————————

481 GCC ATC TGG GTC CTC GCT ACT CTC TTC TCT CTG CCC AAC ACC AGC CTG CAT GGC CTC AGT   540
161 A   I   W   V   L   A   T   L   F   S   L   P   N   T   S   L   H   G   L   S   180

541 CAA CTA ACT GTG CCC TGC CGG GGG CCG GTG CCC GAC TCA GCT ATA TGT TCG CTG GTG GGT   600
181 Q   L   T   V   P   C   R   G   P   V   P   D   S   A   I   C   S   L   V   G   200
                  TM-5————————————————————————————————————————————————————————
601 CCC ATG GAC TTC TAC AAG TTG GTG GTA CTG ACT ACC GCA CTG CTC TTC TTC TGT CTG CCC   660
201 P   M   D   F   Y   K   L   V   V   L   T   T   A   L   L   F   F   C   L   P   220
 ————————————————————————————————————————

661 ATG GTC ACC ATC AGT GTG CTG TAT CTG CTC ATT GGG CTG CGG CTG CGG AGG GAG AGG ATG   720
221 M   V   T   I   S   V   L   Y   L   L   I   G   L   R   L   R   R   E   R   M   240

721 TTG CTC CAA GTG GAG GTC AAG GGC AGG AAA ACC GCA GCA ACC CAG GAG ACC TCC CAC AGA   780
241 L   L   Q   V   E   V   K   G   R   K   T   A   A   T   Q   E   T   S   H   R   260
                                                          TM-6————————————————————
781 AGG ATT CAG CTG CAA GAT AGG GGA CGG AGA CAG GTG ACC AAG ATG CTG TTT GCA CTG GTT   840
261 R   I   Q   L   Q   D   R   G   R   R   Q   V   T   K   M   L   F   A   L   V   280
 ————————————————————————————————————————————————————

841 GTG GTA TTC GGC ATC TGC TGG GCT CCA TTC CAT GCT GAC CGT ATC ATG TGG AGC CTG GTG   900
281 V   V   F   G   I   C   W   A   P   F   H   A   D   R   I   M   W   S   L   V   300
                                                              TM-7————————————————
901 TAT GGA CAC TCA ACG GAA GGC CTG CAC CTG GCC TAC CAG TGT GTC CAC ATT GCC TCT GGC   960
301 Y   G   H   S   T   E   G   L   H   L   A   Y   Q   C   V   H   I   A   S   G   320
 ————————————————————————————————————————————————————————

961 ATC TTC TTC TAT CTC GGC TCA GCA GCC AAC CCG GTG CTC TAC AGC CTC ATG TCT ACT CGC  1020
321 I   F   F   Y   L   G   S   A   A   N   P   V   L   Y   S   L   M   S   T   R   340

1021 TTC CGA GAG ACC TTC CTG CAA GCC CTG GGC CTT GGA ACC CAG TGC TGT CAT CGC CGC CAA  1080
 341 F   R   E   T   F   L   Q   A   L   G   L   G   T   Q   C   C   H   R   R   Q   360

1081 CCC TAT CAT GGC TCC CAT AAC CAC ATC AGG TTG ACC ACA GGC AGC ACC CTG TGT GAC GTG  1140
 361 P   Y   H   G   S   H   N   H   I   R   L   T   T   G   S   T   L   C   D   V   380

1141 GGC CAC AGG AAC AGC AGG GAC GAA CCT CTG GCT GTG AAT GAG GAT CCA GGG TGT CAG CAA  1200
 381 G   H   R   N   S   R   D   E   P   L   A   V   N   E   D   P   G   C   Q   Q   400

1201 GAG ACA GAC CCC TCC TGA                                                        1218
 401 E   T   D   P   S   *                                                          406
```

FIG. 1. DNA and deduced amino acid sequence for mouse (A) and human (B) FM-3. Putative transmembrane α helices are overlined and numbered from 1 to 7. For the murine form, a noncanonical leucine codon is postulated to serve as the initiator.

B

```
 1 ATG GCT TGC AAT GGC AGT GCG GCC AGG GGG CAC TTT GAC CCT GAG GAC TTG AAC CTG ACT  60
 1 M   A   C   N   G   S   A   A   R   G   H   F   D   P   E   D   L   N   L   T   20
                                                                 TM-1
 61 GAC GAG GCA CTG AGA CTC AAG TAC CTG GGG CCC CAG CAG ACA GAG CTG TTC ATG CCC ATC 120
 21 D   E   A   L   R   L   K   Y   L   G   P   Q   Q   T   E   L   F   M   P   I   40

121 TGT GCC ACA TAC CTG CTG ATC TTC GTG GTG GGC GCT GTG GGC AAT GGG CTG ACC TGT CTG 180
 41 C   A   T   Y   L   L   I   F   V   V   G   A   V   G   N   G   L   T   C   L   60
                                                     TM-2
181 GTC ATC CTG CGC CAC AAG GCC ATG CGC ACG CCT ACC AAC TAC TAC CTC TTC AGC CTG GCC 240
 61 V   I   L   R   H   K   A   M   R   T   P   T   N   Y   Y   L   F   S   L   A   80

241 GTG TCG GAC CTG CTG GTG CTG CTG GTG GGC CTG CCC CTG GAG CTC TAT GAG ATG TGG CAC 300
 81 V   S   D   L   L   V   L   L   V   G   L   P   L   E   L   Y   E   M   W   H  100
                                                 TM-3
301 AAC TAC CCC TTC CTG CTG GGC GTT GGT GGC TGC TAT TTC CGC ACG CTA CTG TTT GAG ATG 360
101 N   Y   P   F   L   L   G   V   G   G   C   Y   F   R   T   L   L   F   E   M  120

361 GTC TGC CTG GCC TCA GTG CTC AAC GTC ACT GCC CTG AGC GTG GAA CGC TAT GTG GCC GTG 420
121 V   C   L   A   S   V   L   N   V   T   A   L   S   V   E   R   Y   V   A   V  140
                                                             TM-4
421 GTG CAC CCA CTC CAG GCC AGG TCC ATG GTG ACG CGG GCC CAT GTG CGC CGA GTG CTT GGG 480
141 V   H   P   L   Q   A   R   S   M   V   T   R   A   H   V   R   R   V   L   G  160

481 GCC GTC TGG GGT CTT GCC ATG CTC TGC TCC CTG CCC AAC ACC AGC CTG CAC GGC ATC CGG 540
161 A   V   W   G   L   A   M   L   C   S   L   P   N   T   S   L   H   G   I   R  180

541 CAG CTG CAC GTG CCC TGC CGG GGC CCA GTG CCA GAC TCA GCT GTT TGC ATG CTG GTC CGC 600
181 Q   L   H   V   P   C   R   G   P   V   P   D   S   A   V   C   M   L   V   R  200
                TM-5
601 CCA CGG GCC CTC TAC AAC ATG GTA GTG CAG ACC ACC GCG CTG CTC TTC TTC TGC CTG CCC 660
201 P   R   A   L   Y   N   M   V   V   Q   T   T   A   L   L   F   F   C   L   P  220

661 ATG GCC ATC ATG AGC GTG CTC TAC CTG CTC ATT GGG CTG CGA CTG CGG CGG GAG AGG CTG 720
221 M   A   I   M   S   V   L   Y   L   L   I   G   L   R   L   R   R   E   R   L  240

721 CTG CTC ATG CAG GAG GCC AAG GGC AGG GGC TCT GCA GCA GCC AGG TCC AGA TAC ACC TGC 780
241 L   L   M   Q   E   A   K   G   R   G   S   A   A   A   R   S   R   Y   T   C  260
                                                 TM-6
781 AGG CTC CAG CAG CAC GAT CGG GGC CGG AGA CAA GTG ACC AAG ATG CTG TTT GTC CTG GTC 840
261 R   L   Q   Q   H   D   R   G   R   R   Q   V   T   K   M   L   F   V   L   V  280

841 GTG GTG TTT GGC ATC TGC TGG GCC CCG TTC CAC GCC GAC CGC GTC ATG TGG AGC GTC GTG 900
281 V   V   F   G   I   C   W   A   P   F   H   A   D   R   V   M   W   S   V   V  300
                                                             TM-7
901 TCA CAG TGG ACA GAT GGC CTG CAC CTG GCC TTC CAG CAC GTG CAC GTC ATC TCC GGC ATC 960
301 S   Q   W   T   D   G   L   H   L   A   F   Q   H   V   H   V   I   S   G   I  320

961 TTC TTC TAC CTG GGC TCG GCG GCC AAC CCC GTG CTC TAT AGC CTC ATG TCC AGC CGC TTC 1020
321 F   F   Y   L   G   S   A   A   N   P   V   L   Y   S   L   M   S   S   R   F  340

1021 CGA GAG ACC TTC CAG GAG GCC CTG TGC CTC GGG GCC TGC TGC CAT CGC CTC AGA CCC CGC 1080
 341 R   E   T   F   Q   E   A   L   C   L   G   A   C   C   H   R   L   R   P   R  360

1081 CAC AGC TCC CAC AGC CTC AGC AGG ATG ACC ACA GGC AGC ACC CTG TGT GAT GTG GGC TCC 1140
 361 H   S   S   H   S   L   S   R   M   T   T   G   S   T   L   C   D   V   G   S  380

1141 CTG GGC AGC TGG GTC CAC CCC CTG GCT GGG AAC GAT GGC CCA GAG GCG CAG CAA GAG ACC 1200
 381 L   G   S   W   V   H   P   L   A   G   N   D   G   P   E   A   Q   Q   E   T  400

1201 GAT CCA TCC TGA                                                              1212
 401 D   P   S   *                                                                  404
```

FIG. 1—*Continued*

GenBank databases were monitored daily using the Tblastn program (1) with amino acid sequence from the human GHS-R TM domains 6–7 (residues 265–366). A mouse EST (dEST database Accession No. AA562357, deposited August 18, 1997) derived from a T-cell library was identified with a significant homology score. EST 562357 exhibited good sequence identity (63% DNA, 36% amino acid) to the 3′ end of the gene for the human GHS-R. A murine T cell λXR cDNA library (Stratagene) was screened with the mouse EST 562357 probe (455 bp in length) under high-stringency conditions. Four partial clones were identified after three rounds of screening. Additional clones were isolated from mouse thymus poly(A)$^+$ RNA via 5′ Race Mara-